

An infinite dimensional KAM theorem with application to two dimensional completely resonant beam equation *

Jiansheng Geng, Shidi Zhou

Department of Mathematics, Nanjing University, Nanjing 210093, P.R.China

Email: jgeng@nju.edu.cn; mathsdzhou@126.com

Abstract

In this paper we consider the completely resonant beam equation on \mathbb{T}^2 with cubic nonlinearity on a subspace of $L^2(\mathbb{T}^2)$ which will be explained later. We establish an abstract infinite dimensional KAM theorem and apply it to the completely resonant beam equation. We prove the existence of a class of Whitney smooth small amplitude quasi-periodic solutions corresponding to finite dimensional tori.

Mathematics Subject Classification: Primary 37K55; 35B10

Keywords: KAM theory; Hamiltonian systems; Beam equation; Birkhoff normal form

1 Introduction

In this paper we consider the two dimensional completely resonant beam equation with cubic nonlinearity on a subspace \mathcal{U} of $L^2(\mathbb{T}^2)$:

$$u_{tt} + \Delta_x^2 u + u^3 = 0 \quad u = u(t, x), t \in \mathbb{R}, x \in \mathbb{T}^2 \quad (1.1)$$

Here t is time and x is the spatial variable. The subspace \mathcal{U} is defined by

$$\mathcal{U} = \left\{ u = \sum_{n \in \mathbb{Z}_{odd}^2} u_n \phi_n, \quad \phi_n(x) = e^{i\langle n, x \rangle} \right\} \quad (1.2)$$

where the integer set \mathbb{Z}_{odd}^2 is defined as

$$\mathbb{Z}_{odd}^2 = \{ n = (n_1, n_2) : n_1 \in 2\mathbb{Z} - 1, n_2 \in 2\mathbb{Z} \} \quad (1.3)$$

This idea comes from the work by M.Procesi [29] and we will explain it later in section 2. The solution of "real" completely resonant beam equation (not on \mathbb{Z}_{odd}^2 , just on \mathbb{Z}^2) will be handled in our forthcoming paper.

*This work is partially supported by NSFC grant 11271180.

The infinite dimensional KAM theory with applications to Hamiltonian PDEs has attracted great interests since 1980s. Starting from the remarkable work [6,19,31], a lot of achievements have been made in 1-dimensional Hamiltonian PDEs about the existence of quasi-periodic solutions by the methods of KAM theory. For these work, just refer to [5,12,13,17,18,20-25,32]. But when people turn to the higher dimensional case, the multiplicity of eigenvalues became a great obstacle because it leads to much more complicated small divisor conditions and measure estimates. The first breakthrough comes from Bourgain's work [3] in 1998. In this work, the cumbersome second Melnikov condition is avoided due to the application of the method of multiscale analysis, which essentially is a Nash-Moser iterative procedure instead of Newtonian iteration being widely used in KAM theory. Following this idea, a lot of important work has been made in higher dimensional case (refer to [1,2,4,30]).

However, despite the advantage of avoiding the difficulty of the second Melnikov conditions, there are also drawbacks of multiscale analysis methods. For example, we couldn't see the linear stability of the small-amplitude solutions and it couldn't show us a description of the normal form, which is fundamental in knowing the dynamical structure of an equation. For these reasons, KAM approach is also expected in dealing with higher dimensional equations. The first work comes from Geng and You [14] in 2006, which established the KAM theorem solving higher dimensional beam equations and nonlocal smooth Schrödinger equations with Fourier multiplier. They used the "zero-momentum condition" to avoid the multiplicity of eigenvalues and the regularity property to do the measure estimate. Later in 2010, a remarkable work [8] by Eliasson and Kuksin dealt with quite general case: higher dimensional Schrödinger equations with convolutional type potential and without "zero-momentum condition". To overcome the multiple eigenvalues they studied the distribution of integer points on a sphere and got a normal form with block-diagonal structure, and conducted the measure estimates by developing the technique named "Lipschitz domain". Motivated by their method, the quasi-periodic solutions of completely resonant Schrödinger equation on 2-dimension torus was developed by Geng, Xu and You [11] in 2011, with a very elaborate construction of tangential sites. In this paper, they defines the conception of "Töplitz-Lipschitz" condition and proved that the perturbation satisfies "Töplitz-Lipschitz" condition. Later, in [26,27] C.Procesi and M.Procesi extended this result to higher dimensional case. For other work about higher dimensional equation, just refer to [7,9,10,15,16,28,29].

Let us turn to beam equation now. In [15] Geng and You got the quasi-periodic solutions of beam equation in high dimension with typical constant potential and the nonlinearity is independent on the spatial variable x . Recently, in [7] Eliasson, Grebert and Kuksin got the quasi-periodic solutions of beam equation having typical constant potential in higher dimensional case, and with an elaborate but quite general choice of tangential sites in the sense of probability. they allow that their normal form contain hyperbolic terms which is cumbersome in solving homological equations. Motivated by their work, we want to consider the completely resonant beam equation (1.1). In our case, there are no outer parameters and only the amplitude provides parameters. Compared with the case of typical constant potential, although we have "zero-momentum condition" here, but when doing the normal form before KAM procedure, some terms still couldn't be eliminated because of the loss of outer parameters. We could only get a block-diagonal normal form with finite dimensional block. As a consequence, our normal form is always

related to the angle variable θ , so here the linear stability is not available. Compared with [11], our convenience is that we have regularity property here and needn't verify the complicated "Töeplitz-Lipschitz condition" at each step. But except for this, our normal form structure and KAM iteration is similar to that in [11].

Now we state the choice of tangential sites. Let $S = \{i_j \in \mathbb{Z}_{\text{odd}}^2 : 1 \leq j \leq b\}$ here $b \geq 2$. We say S is admissible if it satisfies the following conditions.

Proposition 1 (*Structure of S*)

- ① Any three of them are not vertices of a rectangle.
- ② For any $n \in \mathbb{Z}_{\text{odd}}^2 \setminus S$, there exists at most one triple $\{i, j, m\}$ with $i, j \in S, m \in \mathbb{Z}_{\text{odd}}^2 \setminus S$ such that

$$\begin{cases} n - m + i - j = 0 \\ |n|^2 - |m|^2 + |i|^2 - |j|^2 = 0 \end{cases}$$

and if it exists, we say (n, m) are resonant in the first type and denote all such n by \mathcal{L}_1 .

- ③ For any $n \in \mathbb{Z}_{\text{odd}}^2 \setminus S$, there exists at most one triple $\{i, j, m\}$ with $i, j \in S, m \in \mathbb{Z}_{\text{odd}}^2 \setminus S$ such that

$$\begin{cases} n + m - i - j = 0 \\ |n|^2 + |m|^2 - |i|^2 - |j|^2 = 0 \end{cases}$$

and if it exists, we say (n, m) are resonant in the second type and denote all such n by \mathcal{L}_2 .

- ④ Any $n \in \mathbb{Z}_{\text{odd}}^2 \setminus S$ shouldn't be in \mathcal{L}_1 and \mathcal{L}_2 at the same time. It means that $\mathcal{L}_1 \cap \mathcal{L}_2 = \emptyset$.

(Here $|\cdot|$ means l^2 norm)

The proof of the existence of admissible sets is postponed in the Appendix, which is a modification of [11].

Now we could state the main theorem.

Theorem 1 Let $S = (i_1, i_2, \dots, i_b) \subseteq \mathbb{Z}_{\text{odd}}^2$ be an admissible set. There exists a Cantor set \mathcal{C} of positive measure, s.t. $\forall \xi = (\xi_1, \xi_2, \dots, \xi_b) \in \mathcal{C}$, equation (1.1) admits a small-amplitude real-valued quasi-periodic solution

$$u(t, x) = \sum_{j=1}^b \sqrt{\xi_j} (e^{i\omega_j t} \phi_{i_j} + e^{-i\omega_j t} \bar{\phi}_{i_j}) + O(|\xi|^{\frac{3}{2}})$$

The outline of this paper is as follows: In section 2 we state some preliminaries and the abstract KAM theorem. In section 3 we deal with the normal form before KAM iteration. In section 4 we conduct one step of KAM iteration: solving homological equation and verifying the new normal form and perturbation. In section 5 we prove uniform convergence and get the invariant torus. In section 6 we complete the measure estimate. The choice of tangential sites is put into the appendix.

2 Preliminaries and statement of the abstract KAM theorem

In this section we introduce some notations and state the abstract KAM theorem which allows the existence of some terms dependent on θ in the normal form part.

To simplify, we only consider the subspace \mathbb{Z}_{odd}^2 (defined in (1.3)) instead of \mathbb{Z}^2 . Given b points $\{i_1, i_2, \dots, i_b\}$ ($b \geq 2$) in \mathbb{Z}_{odd}^2 , denoted by S , which should be an admissible set (defined in Proposition 1), and let \mathbb{Z}_1^2 be the complementary set of S in \mathbb{Z}_{odd}^2 . Denote $z = (z_n)_{n \in \mathbb{Z}_1^2}$ with its conjugate $\bar{z} = (\bar{z}_n)_{n \in \mathbb{Z}_1^2}$. We introduce the weighted norm as follows:

$$\|z\|_{a,\rho} = \sum_{n \in \mathbb{Z}_1^2} |z_n| |n|^a e^{\rho|n|} \quad a, \rho > 0 \quad (2.1)$$

Here $|n| = \sqrt{|n_1|^2 + |n_2|^2}$, $n = (n_1, n_2) \in \mathbb{Z}_1^2$. Denote a neighborhood of $\mathbb{T}^b \times \{I = 0\} \times \{z = 0\} \times \{\bar{z} = 0\}$ by

$$D(r, s) = \{(\theta, I, z, \bar{z}) : |\operatorname{Im} \theta| < r, |I| < s^2, \|z\|_{a,\rho} < s, \|\bar{z}\|_{a,\rho} < s\}$$

Here $|\cdot|$ means the sup-norm of complex vectors.

Let $\alpha = \{\alpha_n\}_{n \in \mathbb{Z}_1^2}, \beta = \{\beta_n\}_{n \in \mathbb{Z}_1^2}$, $\alpha_n, \beta_n \in \mathbb{N}$ with only finitely many non-vanishing components. Denote $z^\alpha \bar{z}^\beta = \prod_{n \in \mathbb{Z}_1^2} z_n^{\alpha_n} \bar{z}_n^{\beta_n}$ and let

$$F(\theta, I, z, \bar{z}) = \sum_{k,l,\alpha,\beta} F_{kl\alpha\beta}(\xi) e^{i\langle k, \theta \rangle} I^l z^\alpha \bar{z}^\beta \quad (2.2)$$

where $\xi \in \mathcal{O} \subseteq \mathbb{R}^b$ is the parameter set. $k = (k_1, \dots, k_b) \in \mathbb{Z}^b$ and $l = (l_1, \dots, l_b) \in \mathbb{N}^b$, $I^l = I_1^{l_1} \dots I_b^{l_b}$. Denote the weighted norm of F by

$$\|F\|_{D(r,s),\mathcal{O}} = \sup_{\xi \in \mathcal{O}, \|z\|_{a,\rho} < s, \|\bar{z}\|_{a,\rho} < s} \sum_{k,l,\alpha,\beta} |F_{kl\alpha\beta}|_{\mathcal{O}} e^{k|r} s^{2|l|} |z^\alpha| |\bar{z}^\beta| \quad (2.3)$$

$$|F_{kl\alpha\beta}|_{\mathcal{O}} = \sup_{\xi \in \mathcal{O}} \sum_{0 \leq d \leq 4} |\partial_\xi^d F_{kl\alpha\beta}| \quad (2.4)$$

where the derivatives with respect to ξ are in the sense of Whitney.

To a function F we define its Hamiltonian vector field by

$$X_F = (F_I, -F_\theta, i\{F_{z_n}\}_{n \in \mathbb{Z}_1^2}, -i\{F_{\bar{z}_n}\}_{n \in \mathbb{Z}_1^2}) \quad (2.5)$$

and the associated weighted norm is

$$\begin{aligned} \|X_F\|_{D(r,s),\mathcal{O}} &:= \|F_I\|_{D(r,s),\mathcal{O}} + \frac{1}{s^2} \|F_\theta\|_{D(r,s),\mathcal{O}} \\ &+ \frac{1}{s} \left(\sum_{n \in \mathbb{Z}_1^2} \|F_{z_n}\|_{D(r,s),\mathcal{O}} |n|^{\bar{a}} e^{|n|\rho} + \sum_{n \in \mathbb{Z}_1^2} \|F_{\bar{z}_n}\|_{D(r,s),\mathcal{O}} |n|^{\bar{a}} e^{|n|\rho} \right) \end{aligned} \quad (2.6)$$

where $\bar{a} > 0$ is a constant and we need $\bar{a} > a$ to measure the regularity property of the perturbation at each iterative step.

The normal form has the following form:

$$\begin{aligned}
H_0 &= N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} \\
N &= \langle \omega(\xi), I \rangle + \sum_{n \in \mathbb{Z}_1^2} \Omega_n z_n \bar{z}_n \\
\mathcal{A} &= \sum_{n \in \mathcal{L}_1} a_n(\xi) e^{i(\theta_i - \theta_j)} z_n \bar{z}_m \\
\mathcal{B} &= \sum_{n \in \mathcal{L}_2} a_n(\xi) e^{-i(\theta_i + \theta_j)} z_n z_m \\
\bar{\mathcal{B}} &= \sum_{n \in \mathcal{L}_2} a_n(\xi) e^{i(\theta_i + \theta_j)} \bar{z}_n \bar{z}_m
\end{aligned}$$

where $\xi \in \mathcal{O}$ is the parameter. For each $n \in \mathcal{L}_1$ or $n \in \mathcal{L}_2$, the 3-triple (m, i, j) is uniquely determined.

For this unperturbed system, it's easy to see that it admits a special solution $(\theta, 0, 0, 0) \rightarrow (\theta + \omega t, 0, 0, 0)$ corresponding to an invariant torus in the phase space. Our goal is to prove that, after removing some parameters, the perturbed system $H = H_0 + P$ still admits invariant torus provided that $\|X_P\|_{D_{a,\rho}(r,s),\mathcal{O}}$ is sufficiently small. To achieve this goal, we require that Hamiltonian H satisfies some conditions:

(A1) Nondegeneracy: The map $\xi \rightarrow \omega(\xi)$ is a C_W^4 diffeomorphism between \mathcal{O} and its image (C_W^4 means C^4 in the sense of Whitney).

(A2) Asymptotics of normal frequencies:

$$\Omega_n = \varepsilon^{-p} |n|^2 + \tilde{\Omega}_n \quad p > 0 \quad (2.7)$$

here $\tilde{\Omega}_n$ is a C_W^4 function of ξ , and $\tilde{\Omega}_n = O(|n|^{-\iota}) \quad \iota > 0$

(A3) Melnikov conditions: Let

$$A_n = \Omega_n \quad n \in \mathbb{Z}_1^2 \setminus (\mathcal{L}_1 \cup \mathcal{L}_2)$$

and

$$\begin{aligned}
A_n &= \begin{pmatrix} \Omega_n + \omega_i & a_n \\ a_m & \Omega_m + \omega_j \end{pmatrix} \quad n \in \mathcal{L}_1 \\
A_n &= \begin{pmatrix} \Omega_n - \omega_i & a_n \\ \bar{a}_m & \Omega_m - \omega_j \end{pmatrix} \quad n \in \mathcal{L}_2
\end{aligned}$$

Then we assume that there exists $\gamma, \tau > 0$, such that

$$\begin{aligned}
|\langle k, \omega \rangle| &\geq \frac{\gamma}{|k|^\tau} \quad k \neq 0 \\
|\det(\langle k, \omega \rangle + A_n)| &\geq \frac{\gamma}{|k|^\tau} \\
|\det(\langle k, \omega \rangle + A_n \otimes I_2 \pm I_2 \otimes A_{n'})| &\geq \frac{\gamma}{|k|^\tau} \quad k \neq 0
\end{aligned}$$

(A4) Boundedness: $\mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P$ is real analytic in each variable θ, I, z, \bar{z} and Whitney smooth in ξ . And we have

$$\|X_{\mathcal{A}}\|_{D_{a,\rho}(r,s),\mathcal{O}} + \|X_{\mathcal{B}}\|_{D_{a,\rho}(r,s),\mathcal{O}} + \|X_{\bar{\mathcal{B}}}\|_{D_{a,\rho}(r,s),\mathcal{O}} < 1, \quad \|X_P\|_{D_{a,\rho}(r,s),\mathcal{O}} < \varepsilon \quad (2.8)$$

(A5) Zero-momentum condition:

The normal form part $\mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P$ satisfy the following condition:

$$\mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P = \sum_{k \in \mathbb{Z}^b, l \in \mathbb{N}^b, \alpha, \beta} (\mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P)_{kl\alpha\beta}(\xi) e^{i\langle k, \theta \rangle} I^l z^\alpha \bar{z}^\beta$$

we have

$$(\mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P)_{kl\alpha\beta} \neq 0 \implies \sum_{j=1}^b k_j i_j + \sum_{n \in \mathbb{Z}_1^2} (\alpha_n - \beta_n) n = 0$$

Now we state our abstract KAM theorem, and as a corollary, we get Theorem 1.

Theorem 2 *Assume that the Hamiltonian $H = N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P$ satisfies condition (A1) – (A5). Let $\gamma > 0$ be sufficiently small, then there exists $\varepsilon > 0$ and $a, \rho > 0$ such that if $\|X_P\|_{D_{a,\rho}(r,s), \mathcal{O}} < \varepsilon$, the following holds: There exists a Cantor subset $\mathcal{O}_\gamma \subseteq \mathcal{O}$ with $\text{meas}(\mathcal{O} \setminus \mathcal{O}_\gamma) = O(\gamma^\varsigma)$ (ς is a positive constant) and two maps which are analytic in θ and C_W^4 in ξ .*

$$\Phi : \mathbb{T}^b \times \mathcal{O}_\gamma \rightarrow D_{a,\rho}(r, s), \quad \tilde{\omega} : \mathcal{O}_\gamma \rightarrow \mathbb{R}^b$$

where Φ is $\frac{\varepsilon}{\gamma^{16}}$ -close to the trivial embedding $\Phi_0 : \mathbb{T}^b \times \mathcal{O} \rightarrow \mathbb{T}^b \times \{0, 0, 0\}$ and $\tilde{\omega}$ is ε -close to the unperturbed frequency ω , such that $\forall \xi \in \mathcal{O}_\gamma$ and $\theta \in \mathbb{T}^b$, the curve $t \rightarrow \Psi(\theta + \tilde{\omega}t, \xi)$ is a quasi-periodic solution of the Hamiltonian equation governed by $H = N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P$.

3 Normal Form

Consider the equation (1.1). The linear operator $-\Delta$ has eigenvalues $\lambda_n = |n|^2$ and corresponding eigenfunctions $\phi_n = \frac{1}{2\pi} e^{i\langle n, x \rangle}$. By scaling $u \rightarrow \varepsilon^{\frac{1}{2}} u$, (1.1) becomes

$$u_{tt} + \Delta^2 u + \varepsilon u^3 = 0 \tag{3.1}$$

Now introduce $v = u_t$ and (3.1) is turned into

$$\begin{aligned} u_t &= v \\ v_t &= -\Delta^2 u - \varepsilon u^3 \end{aligned} \tag{3.2}$$

Let $q = \frac{1}{\sqrt{2}}((-\Delta)^{\frac{1}{2}} u - i(-\Delta)^{-\frac{1}{2}} v)$ and (3.2) becomes

$$-iq_t = -\Delta q + \varepsilon \frac{1}{\sqrt{2}} (-\Delta)^{-\frac{1}{2}} \left((-\Delta)^{-\frac{1}{2}} \left(\frac{q + \bar{q}}{\sqrt{2}} \right) \right)^3 \tag{3.3}$$

Write it in the form of Hamiltonian equation $q_t = i \frac{\partial H}{\partial \bar{q}}$ and we get the Hamiltonian

$$H = \frac{1}{2} \langle -\Delta q, q \rangle + \frac{1}{4} \varepsilon \int_{\mathbb{T}^2} \left((-\Delta)^{-\frac{1}{2}} (q + \bar{q}) \right)^4 dx \tag{3.4}$$

where $\langle \cdot, \cdot \rangle$ is the inner product in $L^2(\mathbb{T}^2)$. Notice that in \mathbb{Z}_{odd}^2 the origin is avoided so $(-\Delta)^{-\frac{1}{2}}$ is well defined. (That is why we use it instead of the whole \mathbb{Z}^2) Now expand q into Fourier series

$$q = \sum_{n \in \mathbb{Z}_{odd}^2} q_n \phi_n \quad (3.5)$$

so the Hamiltonian becomes (justify ε if necessary)

$$\begin{aligned} H = & \sum_{n \in \mathbb{Z}_{odd}^2} \lambda_n |q_n|^2 + \varepsilon \sum_{i+j+k+l=0} \frac{1}{\sqrt{\lambda_i \lambda_j \lambda_k \lambda_l}} (q_i q_j q_k q_l + \bar{q}_i \bar{q}_j \bar{q}_k \bar{q}_l) \\ & + 4\varepsilon \sum_{i+j+k-l=0} \frac{1}{\sqrt{\lambda_i \lambda_j \lambda_k \lambda_l}} (q_i q_j q_k \bar{q}_l + \bar{q}_i \bar{q}_j \bar{q}_k q_l) \\ & + 6\varepsilon \sum_{i+j-k-l=0} \frac{1}{\sqrt{\lambda_i \lambda_j \lambda_k \lambda_l}} (q_i q_j \bar{q}_k \bar{q}_l) \end{aligned} \quad (3.6)$$

Now we state the normal form theorem of H .

Proposition 3.1 *Let S be admissible. For Hamiltonian function (3.6), there exists a symplectic transformation Φ satisfying*

$$H \circ \Phi = \langle \omega, I \rangle + \sum_{n \in \mathbb{Z}_1^2} \Omega_n z_n \bar{z}_n + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P \quad (3.7)$$

where

$$\omega_i = \varepsilon^{-4} \lambda_i + \frac{2}{\lambda_i^2} \xi_i + 4 \sum_{j \in S, j \neq i} \frac{1}{\lambda_i \lambda_j} \xi_j \quad i \in S \quad (3.8)$$

$$\Omega_n = \varepsilon^{-4} \lambda_n + 4 \sum_{j \in S} \frac{1}{\lambda_j \lambda_n} \xi_j \quad n \in \mathbb{Z}_1^2 \quad (3.9)$$

and

$$\mathcal{A} = 4 \sum_{n \in \mathcal{L}_1} \frac{\sqrt{\xi_i \xi_j}}{\sqrt{\lambda_i \lambda_j \lambda_n \lambda_m}} e^{i(\theta_i - \theta_j)} z_n \bar{z}_m \quad (3.10)$$

$$\mathcal{B} = 4 \sum_{n \in \mathcal{L}_2} \frac{\sqrt{\xi_i \xi_j}}{\sqrt{\lambda_i \lambda_j \lambda_n \lambda_m}} e^{i(-\theta_i - \theta_j)} z_n z_m \quad (3.11)$$

$$\bar{\mathcal{B}} = 4 \sum_{n \in \mathcal{L}_2} \frac{\sqrt{\xi_i \xi_j}}{\sqrt{\lambda_i \lambda_j \lambda_n \lambda_m}} e^{i(\theta_i + \theta_j)} \bar{z}_n \bar{z}_m \quad (3.12)$$

$$\begin{aligned} P = & O(\varepsilon^2 |I|^2 + \varepsilon^2 |I| \|z\|_{a,\rho}^2 + \varepsilon |\xi|^{\frac{1}{2}} \|z\|_{a,\rho}^3 + \varepsilon^2 \|z\|_{a,\rho}^4 + \varepsilon^2 |\xi|^3 \\ & + \varepsilon^3 |\xi|^{\frac{5}{2}} \|z\|_{a,\rho} + \varepsilon^4 |\xi|^2 \|z\|_{a,\rho}^2 + \varepsilon^5 |\xi|^{\frac{3}{2}} \|z\|_{a,\rho}^3) \end{aligned} \quad (3.13)$$

Proof : We construct a Hamiltonian function F to induce $\Phi = X_F^1$ which is the time-1 map of F . For convenience, we define three sets as below:

$$\begin{aligned} S_1 = \{(i, j, n, m) : & \quad \textcircled{1} : i - j + n - m = 0 \\ & \quad \textcircled{2} : |i|^2 - |j|^2 + |n|^2 - |m|^2 \neq 0 \\ & \quad \textcircled{3} : \#\{i, j, n, m\} \cap S \geq 2\} \end{aligned} \quad (3.14)$$

and similarly

$$\begin{aligned} S_2 = \{(i, j, n, m) : & \quad \textcircled{1} : i + j + n + m = 0 \\ & \quad \textcircled{2} : |i|^2 + |j|^2 + |n|^2 + |m|^2 \neq 0 \\ & \quad \textcircled{3} : \#\{i, j, n, m\} \cap S \geq 2\} \end{aligned} \quad (3.15)$$

$$\begin{aligned} S_3 = \{(i, j, n, m) : & \quad \textcircled{1} : i + j + n - m = 0 \\ & \quad \textcircled{2} : |i|^2 + |j|^2 + |n|^2 - |m|^2 \neq 0 \\ & \quad \textcircled{3} : \#\{i, j, n, m\} \cap S \geq 2\} \end{aligned} \quad (3.16)$$

we define F as

$$\begin{aligned} F &= \sum_{S_1} \frac{i\varepsilon}{\lambda_i - \lambda_j + \lambda_n - \lambda_m} q_i \bar{q}_j q_n \bar{q}_m \\ &+ \sum_{S_2} \frac{i\varepsilon}{6(\lambda_i + \lambda_j + \lambda_n + \lambda_m)} (q_i q_j q_n q_m - \bar{q}_i \bar{q}_j \bar{q}_n \bar{q}_m) \\ &+ \sum_{S_3} \frac{2i\varepsilon}{3(\lambda_i + \lambda_j + \lambda_n - \lambda_m)} (q_i q_j q_n \bar{q}_m - \bar{q}_i \bar{q}_j \bar{q}_n q_m) \end{aligned} \quad (3.17)$$

(3.6) is put into (set $z_n = q_n, \bar{z}_n = \bar{q}_n, n \notin S$)

$$\begin{aligned} H \circ \Phi &= \sum_{n \in S} \lambda_n |q_n|^2 + \sum_{n \notin S} \lambda_n |z_n|^2 + \varepsilon \sum_{n \in S} \frac{1}{\lambda_n^2} |q_n|^4 \\ &+ 4\varepsilon \sum_{i, j \in S, i \neq j} \frac{1}{\lambda_i \lambda_j} |q_i|^2 |q_j|^2 + 4\varepsilon \sum_{i \in S, n \notin S} \frac{1}{\lambda_i \lambda_n} |q_i|^2 |z_n|^2 \\ &+ 4\varepsilon \sum_{n \in \mathcal{L}_1} \frac{1}{\sqrt{\lambda_i \lambda_j \lambda_n \lambda_m}} q_i \bar{q}_j z_n \bar{z}_m + 4\varepsilon \sum_{n \in \mathcal{L}_2} \frac{1}{\sqrt{\lambda_i \lambda_j \lambda_n \lambda_m}} (q_i q_j \bar{z}_n \bar{z}_m + \bar{q}_i \bar{q}_j z_n z_m) \\ &+ O\left(\varepsilon |q| \|z\|_{a, \rho}^3 + \varepsilon \|z\|_{a, \rho}^4 + \varepsilon^2 |q|^6 + \varepsilon^2 |q|^5 \|z\|_{a, \rho}^3 + \varepsilon^2 |q|^4 \|z\|_{a, \rho}^2 + \varepsilon^2 |q|^3 \|z\|_{a, \rho}^3\right) \end{aligned}$$

Here we need to state a fact: For four points $n, m, i, j \in \mathbb{Z}_{odd}^2$, it could never satisfy $|n|^2 + |m|^2 + |i|^2 - |j|^2 = 0$. If not, we assume $n = (n_1, n_2), m = (m_1, m_2), i = (i_1, i_2), j = (j_1, j_2)$ and in each one the first component is odd and the second component is even. Then we have

$$|n_1|^2 + |m_1|^2 + |i_1|^2 - |j_1|^2 = -(|n_2|^2 + |m_2|^2 + |i_2|^2 - |j_2|^2)$$

The right one can be divided by 4 but the left one couldn't, which is a contradiction. By this fact we know that the set

$$\{(i, j, n, m) \in (\mathbb{Z}_{\text{odd}}^2)^4 : \begin{array}{ll} \textcircled{1} : i + j + n - m = 0 \\ \textcircled{2} : |i|^2 + |j|^2 + |n|^2 - |m|^2 = 0 \end{array}\} \quad (3.18)$$

is empty.

Introduce the action-angle variable in the tangential sites:

$$q_j = \sqrt{I_j + \xi_j} e^{i\theta_j}, \quad \bar{q}_j = \sqrt{I_j + \xi_j} e^{-i\theta_j} \quad (3.19)$$

so we have

$$\begin{aligned} H \circ \Phi &= \sum_{i \in S} \lambda_i (I_i + \xi_i) + \sum_{n \notin S} \lambda_n |z_n|^2 + \varepsilon \sum_{i \in S} \frac{1}{\lambda_i^2} (I_i + \xi_i)^2 \\ &+ 4\varepsilon \sum_{i, j \in S, i \neq j} \frac{1}{\lambda_i \lambda_j} (I_i + \xi_i)(I_j + \xi_j) + 4\varepsilon \sum_{i \in S, n \notin S} \frac{1}{\lambda_i \lambda_n} (I_i + \xi_i) |z_n|^2 \\ &+ 4\varepsilon \sum_{n \in \mathcal{L}_1} \frac{1}{\sqrt{\lambda_i \lambda_j \lambda_n \lambda_m}} \sqrt{(I_i + \xi_i)(I_j + \xi_j)} e^{i(\theta_i - \theta_j)} z_n \bar{z}_m \\ &+ 4\varepsilon \sum_{n \in \mathcal{L}_2} \frac{1}{\sqrt{\lambda_i \lambda_j \lambda_n \lambda_m}} \sqrt{(I_i + \xi_i)(I_j + \xi_j)} e^{i(-\theta_i - \theta_j)} z_n z_m \\ &+ 4\varepsilon \sum_{n \in \mathcal{L}_2} \frac{1}{\sqrt{\lambda_i \lambda_j \lambda_n \lambda_m}} \sqrt{(I_i + \xi_i)(I_j + \xi_j)} e^{i(\theta_i + \theta_j)} \bar{z}_n \bar{z}_m \\ &+ O(\varepsilon |\xi|^{\frac{1}{2}} \|z\|_{a, \rho}^3 + \varepsilon \|z\|_{a, \rho}^4 + \varepsilon^2 |\xi|^3 + \varepsilon^2 |\xi|^{\frac{5}{2}} \|z\|_{a, \rho} + \varepsilon^2 |\xi|^2 \|z\|_{a, \rho}^2 + \varepsilon^2 |\xi|^{\frac{3}{2}} \|z\|_{a, \rho}^3) \\ &= \sum_{i \in S} (\lambda_i + \frac{2\varepsilon}{\lambda_i^2} \xi_i + 4\varepsilon \sum_{j \in S, i \neq j} \frac{1}{\lambda_i \lambda_j} \xi_j) \omega_i \\ &+ \sum_{n \notin S} (\lambda_n + 4\varepsilon \sum_{i \in S} \frac{1}{\lambda_i \lambda_n} \xi_i) |z_n|^2 \\ &+ 4\varepsilon \sum_{n \in \mathcal{L}_1} \frac{1}{\sqrt{\lambda_i \lambda_j \lambda_n \lambda_m}} \sqrt{\xi_i \xi_j} e^{i(\theta_i - \theta_j)} z_n \bar{z}_m \\ &+ 4\varepsilon \sum_{n \in \mathcal{L}_2} \frac{1}{\sqrt{\lambda_i \lambda_j \lambda_n \lambda_m}} \sqrt{\xi_i \xi_j} e^{i(-\theta_i - \theta_j)} z_n z_m \\ &+ 4\varepsilon \sum_{n \in \mathcal{L}_2} \frac{1}{\sqrt{\lambda_i \lambda_j \lambda_n \lambda_m}} \sqrt{\xi_i \xi_j} e^{i(\theta_i + \theta_j)} \bar{z}_n \bar{z}_m \\ &+ O(\varepsilon |I|^2 + \varepsilon |I| \|z\|_{a, \rho}^2 + \varepsilon |\xi|^{\frac{1}{2}} \|z\|_{a, \rho}^3 + \varepsilon \|z\|_{a, \rho}^4 + \varepsilon^2 |\xi|^3 \\ &+ \varepsilon^2 |\xi|^{\frac{5}{2}} \|z\|_{a, \rho} + \varepsilon^2 |\xi|^2 \|z\|_{a, \rho}^2 + \varepsilon^2 |\xi|^{\frac{3}{2}} \|z\|_{a, \rho}^3) \end{aligned}$$

By scaling in variables:

$$\xi \rightarrow \varepsilon^3 \xi, \quad I \rightarrow \varepsilon^5 I, \quad z \rightarrow \varepsilon^{\frac{5}{2}} z, \quad \bar{z} \rightarrow \varepsilon^{\frac{5}{2}} \bar{z}$$

and scale time $t \rightarrow \varepsilon^9 t$ we get the Hamiltonian function as follows:

$$H = \langle \omega, I \rangle + \langle \Omega z, z \rangle + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P \quad (3.20)$$

where

$$\omega_i = \varepsilon^{-4} \lambda_i + \frac{2}{\lambda_i^2} \xi_i + 4 \sum_{j \in S, j \neq i} \frac{1}{\lambda_i \lambda_j} \xi_j \quad (3.21)$$

$$\Omega_n = \varepsilon^{-4} \lambda_n + 4 \sum_{j \in S} \frac{1}{\lambda_j \lambda_n} \xi_j \quad (3.22)$$

$$\mathcal{A} = 4 \sum_{n \in \mathcal{L}_1} \frac{\sqrt{\xi_i \xi_j}}{\sqrt{\lambda_i \lambda_j \lambda_n \lambda_m}} e^{i(\theta_i - \theta_j)} z_n \bar{z}_m \quad (3.23)$$

$$\mathcal{B} = 4 \sum_{n \in \mathcal{L}_2} \frac{\sqrt{\xi_i \xi_j}}{\sqrt{\lambda_i \lambda_j \lambda_n \lambda_m}} e^{i(-\theta_i - \theta_j)} z_n z_m \quad (3.24)$$

$$\bar{\mathcal{B}} = 4 \sum_{n \in \mathcal{L}_2} \frac{\sqrt{\xi_i \xi_j}}{\sqrt{\lambda_i \lambda_j \lambda_n \lambda_m}} e^{i(\theta_i + \theta_j)} \bar{z}_n \bar{z}_m \quad (3.25)$$

$$P = O(\varepsilon^2 |I|^2 + \varepsilon^2 |I| \|z\|_{a,\rho}^2 + \varepsilon |\xi|^{\frac{1}{2}} \|z\|_{a,\rho}^3 + \varepsilon^2 \|z\|_{a,\rho}^4 + \varepsilon^2 |\xi|^3 + \varepsilon^3 |\xi|^{\frac{5}{2}} \|z\|_{a,\rho} \\ + \varepsilon^4 |\xi|^2 \|z\|_{a,\rho}^2 + \varepsilon^5 |\xi|^{\frac{3}{2}} \|z\|_{a,\rho}^3) \quad (3.26)$$

Now we verify that the normal form (3.7) – (3.13) satisfy condition **(A1)** – **(A5)**.

Verifying(A1): By (3.8) we get

$$\frac{\partial \omega}{\partial \xi} = (a_{ij})_{1 \leq i, j \leq b} \quad (3.27)$$

where $a_{ij} = \frac{2}{\lambda_i^2}$ if $i = j$ and $a_{ij} = \frac{4}{\lambda_i \lambda_j}$ if $i \neq j$. It's easy to see that this matrix is non-degenerate.

Verifying(A2): By (3.9), just take $p = 4, \iota = 2$.

Verifying(A3): Recall the definition in condition (A3), we only verify the most complicated case:

$$\det(\langle k, \omega \rangle + A_n \otimes I_2 - I_2 \otimes A_{n'}) \quad (3.28)$$

where $n, n' \in \mathcal{L}_1 \cup \mathcal{L}_2$. We verify two facts: (3.28) is a polynomial of parameter ξ with degree 4 and it couldn't be equivalently zero. For the former one, notice that $\lambda I + A \otimes I_2 - I_2 \otimes B = (\lambda I + A) \otimes I - I \otimes B$ (here $|\cdot|$ means determinant) and using the formula

$$|A \otimes I \pm I \otimes B| = (|A| - |B|)^2 + |A|(tr(B))^2 + |B|(tr(A))^2 \pm (|A| + |B|)tr(A)tr(B)$$

then we get it. For the latter one, it's the same as that in [11]. By this, we could get

$$|\partial_\xi^4 (\det(\langle k, \omega \rangle + A_n \otimes I_2 \pm I_2 \otimes A_{n'}))| > c|k|$$

So by excluding parameters with measure $O(\gamma^{\frac{1}{4}})$, we have

$$|(\det(\langle k, \omega \rangle + A_n \otimes I_2 \pm I_2 \otimes A_{n'}))| > \frac{\gamma}{|k|^\tau} \quad k \neq 0$$

For the verification of **(A4)** and **(A5)**, just refer to [14].

4 KAM Iteration

We prove Theorem 2 by a KAM iteration which involves an infinite sequence of change of variables. Each step of KAM iteration makes the perturbation smaller than the previous step at the cost of excluding a small set of parameters. We have to prove the convergence of the iteration and estimate the measure of the excluded set after infinite KAM steps.

At the ν -step of the KAM iteration, we consider a Hamiltonian vector field with

$$H_\nu = N_\nu + \mathcal{A}_\nu + \mathcal{B}_\nu + \bar{\mathcal{B}}_\nu + P_\nu = \langle \omega_\nu, I \rangle + \sum_{n \in \mathbb{Z}_1^2} \Omega_n^\nu z_n \bar{z}_n + \mathcal{A}_\nu + \mathcal{B}_\nu + \bar{\mathcal{B}}_\nu + P_\nu$$

where $A_\nu + B_\nu + \bar{B}_\nu + P_\nu$ is defined in $D(r_\nu, s_\nu) \times \mathcal{O}_{\nu-1}$

We construct a map

$$\Phi_\nu : D(r_{\nu+1}, s_{\nu+1}) \times \mathcal{O}_\nu \rightarrow D(r_\nu, s_\nu) \times \mathcal{O}_{\nu-1}$$

so that the vector field $X_{H_\nu \circ \Phi_\nu}$ defined on $D(r_{\nu+1}, s_{\nu+1})$ satisfies

$$\|X_{P_{\nu+1}}\|_{D(r_{\nu+1}, s_{\nu+1}), \mathcal{O}_\nu} = \|X_{H_\nu \circ \Phi_\nu} - X_{N_{\nu+1} + \mathcal{A}_{\nu+1} + \mathcal{B}_{\nu+1} + \bar{\mathcal{B}}_{\nu+1}}\|_{D(r_{\nu+1}, s_{\nu+1}), \mathcal{O}_\nu} \leq \varepsilon_\nu^\kappa, \quad \kappa > 1$$

and the new Hamiltonian still satisfies (A1) – (A5).

To simplify notations, in the following text, the quantities without subscripts refer to quantities at the ν th step, while the quantities with subscripts + denote the corresponding quantities at the $(\nu + 1)$ th step. Let's consider the Hamiltonian defined in $D(r, s) \times \mathcal{O}$:

$$\begin{aligned} H &= N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P \\ &= e + \langle \omega(\xi), I \rangle + \sum_{n \in \mathbb{Z}_1^2} \Omega_n z_n \bar{z}_n + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P(\theta, I, z, \bar{z}, \xi, \varepsilon) \end{aligned} \quad (4.1)$$

We assume that for $\xi \in \mathcal{O}$, one has

$$\begin{aligned} |\langle k, \omega \rangle| &\geq \frac{\gamma}{|k|^\tau} \quad k \neq 0 \\ |\det(\langle k, \omega \rangle I + A_n)| &\geq \frac{\gamma}{|k|^\tau} \\ |\det(\langle k, \omega \rangle I \pm A_n \otimes I_2 \pm I_2 \otimes A_{n'})| &\geq \frac{\gamma}{|k|^\tau} \quad k \neq 0 \end{aligned} \quad (4.2)$$

where $A_n = \Omega_n$ for $n \in \mathbb{Z}_1^2 \setminus (\mathcal{L}_1 \cup \mathcal{L}_2)$ and

$$\begin{aligned} A_n &= \begin{pmatrix} \Omega_n + \omega_i & a_n \\ a_m & \Omega_m + \omega_j \end{pmatrix} \quad n \in \mathcal{L}_1 \\ A_n &= \begin{pmatrix} \Omega_n - \omega_i & a_n \\ \bar{a}_m & \Omega_m - \omega_j \end{pmatrix} \quad n \in \mathcal{L}_2 \end{aligned}$$

where (n, m) are resonant pairs and (i, j) is uniquely determined by (n, m) .

Expand P into Fourier-Taylor series $P = \sum_{k,l,\alpha,\beta} P_{kl\alpha\beta} I^l e^{i\langle k,\theta \rangle} z^\alpha \bar{z}^\beta$ and by (A5) we get that

$$P_{kl\alpha\beta} = 0 \quad \text{if} \quad \sum_{1 \leq j \leq b} k_j i_j + \sum_{n \in \mathbb{Z}_1^2} (\alpha_n - \beta_n) n \neq 0 \quad (4.3)$$

We now let $0 < r_+ < r$ and define

$$s_+ = \frac{1}{4} s \varepsilon^{\frac{1}{3}}, \quad \varepsilon_+ = c \gamma^{-16} (r - r_+)^{-c} \varepsilon^{\frac{4}{3}} \quad (4.4)$$

Here and later, the letter c denotes suitable (possible different) constants independent of the iteration steps.

Now we describe how to construct a set $\mathcal{O}_+ \subseteq \mathcal{O}$ and a change of variables $\Phi : D_+ \times \mathcal{O}_+ = D(r_+, s_+) \times \mathcal{O}_+ \rightarrow D(r, s) \times \mathcal{O}$ such that the transformed Hamiltonian $H_+ = H \circ \Phi = N + \mathcal{A}_+ + \mathcal{B}_+ + \bar{\mathcal{B}}_+ + P_+$ satisfies assumptions (A1) – (A5) with new parameters ε_+, r_+, s_+ and with $\xi \in \mathcal{O}_+$.

4.1 Homological Equation

Expand P into Fourier-Taylor series

$$P = \sum_{k,l,\alpha,\beta} P_{kl\alpha\beta} I^l e^{i\langle k,\theta \rangle} z^\alpha \bar{z}^\beta \quad (4.5)$$

where $k \in \mathbb{Z}^b, l \in \mathbb{N}^b$ and the multi-indices α, β run over the set of all infinite dimensional vectors $\alpha = (\cdots, \alpha_n, \cdots)_{n \in \mathbb{Z}_1^2}, \beta = (\cdots, \beta_n, \cdots)_{n \in \mathbb{Z}_1^2}$ with finitely many nonzero components of positive integers. And by (A5) we get that

$$P_{kl\alpha\beta} = 0 \quad \text{if} \quad \sum_{1 \leq j \leq b} k_j i_j + \sum_{n \in \mathbb{Z}_1^2} (\alpha_n - \beta_n) n \neq 0 \quad (4.6)$$

Consider its quadratic truncation R :

$$\begin{aligned} R(\theta, I, z, \bar{z}) &= R_0 + R_1 + R_2 \\ &= \sum_{k, |l| \leq 1} P_{kl00} e^{i\langle k,\theta \rangle} I^l \\ &\quad + \sum_{k,n} (P_n^{k10} z_n + P_n^{k01} \bar{z}_n) e^{i\langle k,\theta \rangle} \\ &\quad + \sum_{k,n,m} (P_{nm}^{k20} z_n z_m + P_{nm}^{k11} z_n \bar{z}_m + P_{nm}^{k02} \bar{z}_n \bar{z}_m) e^{i\langle k,\theta \rangle} \end{aligned} \quad (4.7)$$

where $P_n^{k10} = P_{kl\alpha\beta}$ with $\alpha = e_n, \beta = 0$, $P_n^{k01} = P_{kl\alpha\beta}$ with $\alpha = 0, \beta = e_n$, here e_n denotes the vector with the n th component being 1 and the other components being zero. Similarly, $P_{nm}^{k20} = P_{kl\alpha\beta}$ with $\alpha = e_n + e_m, \beta = 0$; $P_{nm}^{k11} = P_{kl\alpha\beta}$ with $\alpha = e_n, \beta = e_m$; $P_{nm}^{k02} = P_{kl\alpha\beta}$ with $\alpha = 0, \beta = e_n + e_m$.

Rewrite H as $H = N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + R + (P - R)$. Due to the choice of $s_+ \ll s$ and the definition of the norm, it follows immediately

$$\|X_R\|_{D(r,s),\mathcal{O}} \leq \|X_P\|_{D(r,s),\mathcal{O}} \leq \varepsilon \quad (4.8)$$

and in $D(r, s_+)$

$$\|X_{P-R}\|_{D(r, s_+)} \leq \varepsilon_+ \quad (4.9)$$

In the following, we will construct a Hamiltonian function F satisfying (A5) and with the same form of R defined in $D_+ = D(r_+, s_+)$ such that the time one map X_F^1 of the Hamiltonian vector field X_F defines a map from D_+ to D and puts H into H_+ . Precisely, one has

$$\begin{aligned} H \circ X_F^1 &= (N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + R) \circ X_F^1 + (P - R) \circ X_F^1 \\ &= N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} \\ &+ \{N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}}, F\} + R \end{aligned} \quad (4.10)$$

$$+ \int_0^1 (1-t) \{ \{N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}}, F\}, F \} \circ X_F^t dt \quad (4.11)$$

$$+ \int_0^1 \{R, F\} \circ X_F^t dt + (P - R) \circ X_F^1 \quad (4.12)$$

So we get the linearized homological equation:

$$\{N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}}, F\} + R = \hat{N} + \hat{\mathcal{A}} + \hat{\mathcal{B}} + \hat{\bar{\mathcal{B}}} \quad (4.13)$$

where

$$\hat{N} = P_{0000} + \langle \hat{\omega}, I \rangle + \sum_{n \in \mathbb{Z}_1^2} P_{nn}^{011} z_n \bar{z}_n \quad \hat{\omega} = (P_{0100})_{|l|=1} \quad (4.14)$$

$$\hat{\mathcal{A}} = \sum_{n \in \mathcal{L}_1} P_{nm}^{11} e^{i(\theta_i - \theta_j)} z_n \bar{z}_m \quad (4.15)$$

$$\hat{\mathcal{B}} = \sum_{n \in \mathcal{L}_2} P_{nm}^{20} e^{-i(\theta_i + \theta_j)} z_n z_m \quad (4.16)$$

$$\hat{\bar{\mathcal{B}}} = \sum_{n \in \mathcal{L}_2} P_{nm}^{02} e^{i(\theta_i + \theta_j)} \bar{z}_n \bar{z}_m \quad (4.17)$$

We define $N_+ = N + \hat{N}$, $\mathcal{A}_+ = \mathcal{A} + \hat{\mathcal{A}}$, $\mathcal{B}_+ = \mathcal{B} + \hat{\mathcal{B}}$, $\bar{\mathcal{B}}_+ = \bar{\mathcal{B}} + \hat{\bar{\mathcal{B}}}$ and

$$\begin{aligned} P_+ &= \int_0^1 (1-t) \{ \{N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}}, F\}, F \} \circ X_F^t dt \\ &+ \int_0^1 \{R, F\} \circ X_F^t dt + (P - R) \circ X_F^1 \end{aligned} \quad (4.18)$$

We construct the Hamiltonian function F as below:

$$\begin{aligned} F(\theta, I, z, \bar{z}) &= F^0 + F^1 + F^{10} + F^{01} + F^{20} + F^{11} + F^{02} \\ &= F^0(\theta) + \langle F^1(\theta), I \rangle + \langle F^{10}(\theta), z \rangle + \langle F^{01}(\theta), \bar{z} \rangle \\ &+ \langle F^{20}(\theta) z, z \rangle + \langle F^{11}(\theta) z, \bar{z} \rangle + \langle F^{02}(\theta) \bar{z}, \bar{z} \rangle \end{aligned} \quad (4.19)$$

Now (4.13) turns to

$$\{N, F^0 + F^1\} + R^0 + R^1 - P_{0000} - \langle \hat{\omega}, I \rangle = 0 \quad (4.20)$$

$$\{N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}}, F^{10} + F^{01}\} + R^{10} + R^{01} = 0 \quad (4.21)$$

and the most complicated

$$\begin{aligned} & \{N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}}, F^{20} + F^{11} + F^{02}\} + R^{20} + R^{11} + R^{02} \\ &= \sum_{n \in \mathbb{Z}_1^2} P_{nn}^{011} z_n \bar{z}_n + \hat{\mathcal{A}} + \hat{\mathcal{B}} + \hat{\bar{\mathcal{B}}} \end{aligned} \quad (4.22)$$

Solving(4.20): $F^j(\theta) = \sum_{k \neq 0} F_k^j e^{i\langle k, \theta \rangle}$, $j = 0, 1$. By comparing the Fourier coefficients we get

$$F_k^j = \frac{-i}{\langle k, \omega \rangle} P_k^j, \quad j = 0, 1, k \neq 0$$

and according to assumption (4.2) we get

$$|\langle k, \omega(\xi) \rangle| \geq \frac{\gamma}{|k|^\tau}, k \neq 0, \xi \in \mathcal{O}$$

so one has the estimate

$$|F_k^j|_{\mathcal{O}} \leq \gamma^{-16} k^{16\tau+16} |P_k^j|_{\mathcal{O}} \quad (4.23)$$

Solving(4.21): We decompose this part into three cases:

(1) If $n \in \mathbb{Z}_1^2 \setminus \{\mathcal{L}_1 \cup \mathcal{L}_2\}$, one has

$$\begin{aligned} (\langle k, \omega \rangle + \Omega_n) F_{k,n}^{10} &= -i R_{k,n}^{10} \\ (\langle k, \omega \rangle - \Omega_n) F_{k,n}^{10} &= -i R_{k,n}^{01} \end{aligned} \quad (4.24)$$

(2) If $n \in \mathcal{L}_1$ and the corresponding resonant group m, i, j , one has

$$\begin{aligned} (\langle k + e_i, \omega \rangle + \Omega_n) F_{k+e_i,n}^{10} + a_n F_{k+e_j,m}^{10} &= -i R_{k+e_i,n}^{10} \\ (\langle k + e_j, \omega \rangle + \Omega_m) F_{k+e_j,m}^{10} + a_m F_{k+e_i,n}^{10} &= -i R_{k+e_j,m}^{10} \end{aligned} \quad (4.25)$$

(3) If $n \in \mathcal{L}_2$, one has

$$\begin{aligned} (\langle k - e_i, \omega \rangle + \Omega_n) F_{k-e_i,n}^{10} - a_n F_{k+e_j,m}^{01} &= -i R_{k-e_i,n}^{10} \\ (\langle k + e_j, \omega \rangle - \Omega_m) F_{k+e_j,m}^{01} + \bar{a}_m F_{k-e_i,n}^{10} &= -i R_{k+e_j,m}^{01} \end{aligned} \quad (4.26)$$

The above three equations have the coefficient matrix of the form $\langle k, \omega \rangle I + A_n$ and by the assumption (4.2) we know that $|\det(\langle k, \omega \rangle I + A_n)| \geq \frac{\gamma}{|k|^\tau}$. So we get the estimate

(1): If $n \in \mathbb{Z}_1^2 \setminus \{\mathcal{L}_1 \cup \mathcal{L}_2\}$, one has

$$|F_{k,n}^{10}|_{\mathcal{O}}, |F_{k,n}^{01}|_{\mathcal{O}} \leq c \gamma^{-16} k^{16\tau+16} \max\{|R_{k,n}^{10}|_{\mathcal{O}}, |R_{k,n}^{01}|_{\mathcal{O}}\} \quad (4.27)$$

(2): If $n \in \mathcal{L}_1$ and the corresponding resonant group m, i, j , one has

$$|F_{k+e_i,n}^{10}|_{\mathcal{O}}, |F_{k+e_j,m}^{10}|_{\mathcal{O}} \leq c \gamma^{-16} k^{16\tau+16} \max\{|R_{k+e_i,n}^{10}|_{\mathcal{O}}, |R_{k+e_j,m}^{10}|_{\mathcal{O}}\} \quad (4.28)$$

(3) If $n \in \mathcal{L}_2$, one has

$$|F_{k-e_i,n}^{10}|_{\mathcal{O}}, |F_{k+e_j,m}^{01}|_{\mathcal{O}} \leq c \gamma^{-16} k^{16\tau+16} \max\{|R_{k-e_i,n}^{10}|_{\mathcal{O}}, |R_{k+e_j,m}^{01}|_{\mathcal{O}}\} \quad (4.29)$$

Solving(4.22): Similarly, we also decompose them into three parts. In this case, the coefficient matrix has the form of

$$\langle k, \omega \rangle \pm A_n \otimes I \pm I \otimes A_{n'} \quad n, n' \in \mathbb{Z}_1^2$$

By the assumption:

$$|\det(\langle k, \omega \rangle \pm A_n \otimes I \pm I \otimes A_{n'})| \geq \frac{\gamma}{|k|^\tau} \quad k \neq 0$$

(1): If $n, n' \in \mathbb{Z}_1^2 \setminus (\mathcal{L}_1 \cup \mathcal{L}_2)$, one has

$$\begin{aligned} (\langle k, \omega \rangle + \Omega_n + \Omega_{n'}) F_{k,nn'}^{20} &= -i R_{k,nn'}^{20} \\ (\langle k, \omega \rangle + \Omega_n - \Omega_{n'}) F_{k,nn'}^{11} &= -i R_{k,nn'}^{11} \\ (\langle k, \omega \rangle - \Omega_n - \Omega_{n'}) F_{k,nn'}^{02} &= -i R_{k,nn'}^{02} \end{aligned}$$

and we get the estimate

$$|F_{k,nn'}^{20}|, |F_{k,nn'}^{11}|, |F_{k,nn'}^{02}| \leq c\gamma^{-16} k^{16\tau+16} \max\{|R_{k,nn'}^{20}|, |R_{k,nn'}^{11}|, |R_{k,nn'}^{02}|\} \quad (4.30)$$

(2): If $n \in \mathbb{Z}_1^2 \setminus \{\mathcal{L}_1 \cup \mathcal{L}_2\}$, $n' \in \mathcal{L}_1$, one has

$$\begin{aligned} (\langle k - e_{i'}, \omega \rangle + \Omega_n - \Omega_{n'}) F_{k-e_{i'},nn'}^{11} - a_{n'} F_{k-e_{i'},nm'}^{11} &= -i R_{k-e_{i'},nn'}^{11} \\ (\langle k - e_{j'}, \omega \rangle + \Omega_n - \Omega_{m'}) F_{k-e_{j'},nm'}^{11} - a_{m'} F_{k-e_{j'},nn'}^{11} &= -i R_{k-e_{j'},nm'}^{11} \end{aligned}$$

we have the estimate

$$|F_{k-e_{i'},nn'}^{11}|_{\mathcal{O}}, |F_{k-e_{j'},nm'}^{11}|_{\mathcal{O}} \leq c\gamma^{-16} k^{16\tau+16} \max\{|R_{k-e_{i'},nn'}^{11}|_{\mathcal{O}}, |R_{k-e_{j'},nm'}^{11}|_{\mathcal{O}}\} \quad (4.31)$$

when $n' \in \mathcal{L}_2$ is similar, the estimate is similar.

(3): If $n \in \mathcal{L}_1, n' \in \mathcal{L}_2$, one has

$$\begin{aligned} (\langle k + e_i + e_{i'}, \omega \rangle + \Omega_n - \Omega_{n'}) F_{k+e_i+e_{i'},nn'}^{11} + a_{n'} F_{k+e_i-e_{j'},nm'}^{20} + a_n F_{k+e_j+e_{i'},mn'}^{11} &= -i R_{k+e_i+e_{i'},nn'}^{11} \\ (\langle k + e_i - e_{j'}, \omega \rangle + \Omega_n + \Omega_{m'}) F_{k+e_i-e_{j'},nm'}^{20} - \bar{a}_{m'} F_{k+e_i+e_{i'},nn'}^{11} + a_n F_{k+e_j-e_{j'},mm'}^{20} &= -i R_{k+e_i-e_{j'},nm'}^{20} \\ (\langle k + e_j + e_{i'}, \omega \rangle + \Omega_m - \Omega_{n'}) F_{k+e_j+e_{i'},mn'}^{11} + a_m F_{k+e_i+e_{i'},nn'}^{11} + a_n F_{k+e_i-e_{j'},mm'}^{20} &= -i R_{k+e_j+e_{i'},mn'}^{11} \\ (\langle k + e_j - e_{j'}, \omega \rangle + \Omega_m + \Omega_{m'}) F_{k+e_j-e_{j'},mm'}^{20} + a_m F_{k+e_i-e_{j'},nm'}^{20} - \bar{a}_{m'} F_{k+e_j+e_{i'},mn'}^{11} &= -i R_{k+e_j-e_{j'},mm'}^{20} \end{aligned}$$

and we get the estimate

$$\begin{aligned} &|F_{k+e_i+e_{i'},nn'}^{11}|_{\mathcal{O}}, |F_{k+e_i-e_{j'},nm'}^{20}|_{\mathcal{O}}, |F_{k+e_j+e_{i'},mn'}^{11}|_{\mathcal{O}}, |F_{k+e_j-e_{j'},mm'}^{20}|_{\mathcal{O}} \\ &\leq c\gamma^{-16} |k|^{16\tau+16} \max\{|R_{k+e_i+e_{i'},nn'}^{11}|_{\mathcal{O}}, |R_{k+e_i-e_{j'},nm'}^{20}|_{\mathcal{O}}, |R_{k+e_j+e_{i'},mn'}^{11}|_{\mathcal{O}}, |R_{k+e_j-e_{j'},mm'}^{20}|_{\mathcal{O}}\} \end{aligned} \quad (4.32)$$

when $n \in \mathcal{L}_1, n' \in \mathcal{L}_1$ or $n \in \mathcal{L}_2, n' \in \mathcal{L}_2$, the estimate is similar.

Now we could give the small-divisor condition in the next step with new parameters. For simplicity, we only consider the most complicated case: the second Melnikov condition. Assume that

$$|\det(\langle k, \omega \rangle + A_n \otimes I_2 - I_2 \otimes A_{n'})| > \frac{\gamma}{|k|^\tau}$$

then we have

$$\begin{aligned}
& |\det(\langle k, \omega_+ \rangle + A_n^+ \otimes I_2 - I_2 \otimes A_{n'}^+)| \\
& > |\det(\langle k, \omega \rangle + A_n \otimes I_2 - I_2 \otimes A_{n'})| - c(|k||\hat{\omega}| + \max\{|\hat{a}_n|, |\hat{a}_{n'}|, |\hat{\Omega}_n|, |\hat{\Omega}_{n'}|\}) \\
& > \frac{\gamma}{|k|^\tau} - c|k|\varepsilon > \frac{\gamma_+}{|k|^\tau}
\end{aligned}$$

provided $|k| < K$ where $K = c(\frac{\gamma - \gamma_+}{\varepsilon})^{\frac{1}{\tau+1}}$. So the small divisor condition in the next step holds automatically for $|k| < K$ and we will deal with other terms in section 6.

4.2 Estimation of coordinate transformation and new perturbation

With the similar methods in [14], we could get the estimates of X_F and ϕ_F^t , just with different parameters.

Lemma 4.1 *Let $D_i = D(r_+ + \frac{i}{4}(r - r_+), \frac{i}{4}s)$, $0 \leq i \leq 4$. Then we get*

$$\|X_F\|_{D_{3,\mathcal{O}}} \leq c\gamma^{-16}(r - r_+)^{-c}\varepsilon \quad (4.33)$$

Lemma 4.2 *Let $\eta = \varepsilon^{\frac{1}{3}}$, $D_{i\eta} = D(r_+ + \frac{1}{4}(r - r_+), \frac{i}{4}\eta s)$, $0 \leq i \leq 4$. If $\varepsilon \ll \gamma^{16}(r - r_+)^c$, then we have*

$$\phi_F^t : D_{2\eta} \rightarrow D_{3\eta}, \quad -1 \leq t \leq 1 \quad (4.34)$$

and

$$\|D\phi_F^t - Id\|_{D_{1\eta}} \leq c\gamma^{-16}(r - r_+)^{-c}\varepsilon \quad (4.35)$$

With above estimates, we could give the estimate of new perturbations. We have

$$P_+ = \int_0^1 \{R(t), F\} \circ \phi_F^t dt + (P - R) \circ \phi_F^1$$

where $R(t) = R + (1 - t)\{N, F\} = (1 - t)(N_+ - N) + tR$ and

$$X_{P_+} = \int_0^1 (\phi_F^t)^* X_{\{R(t), F\}} dt + (\phi_F^1)^* X_{(P-R)}$$

By Lemma 4.1, we get

$$\|D\phi_F^t\|_{D_{1\eta}} \leq 1 + \|D\phi_F^t - Id\|_{D_{1\eta}} \leq 2$$

At the same time, by Cauchy estimate, one has

$$\|X_{\{R(t), F\}}\|_{D_{2\eta}} \leq c\gamma^{-16}(r - r_+)^{-c}\eta^{-2}\varepsilon^2$$

One the other hand, we have

$$\|X_{(P-R)}\|_{D_{2\eta}} \leq c\eta\varepsilon$$

To sum up, P_+ is bounded by

$$\|X_{P_+}\|_{D(r_+, s_+)} \leq c\eta\varepsilon + c\gamma^{-16}(r - r_+)^{-c}\eta^{-2}\varepsilon^2 \leq c\varepsilon_+$$

5 Iterative Lemma and Convergence

For fixed parameters $r, s, \varepsilon, \gamma$, at the ν th step of the iterative procedure, we define the sequence

$$\begin{aligned}
r_\nu &= r(1 - \sum_{i=2}^{\nu+1} 2^{-i}) \\
\varepsilon_\nu &= c\gamma^{-16}(r_{\nu-1} - r_\nu)^{-c}\varepsilon_{\nu-1}^{\frac{4}{3}} \\
\gamma_\nu &= \gamma(1 - \sum_{i=2}^{\nu+1} 2^{-i}) \\
\eta_\nu &= \varepsilon_\nu^{\frac{1}{3}} \\
s_\nu &= \frac{1}{4}\eta_{\nu-1}s_{\nu-1} \\
K_\nu &= c(\varepsilon_\nu^{-1}(\gamma_\nu - \gamma_{\nu+1}))^{\frac{1}{\tau+1}}
\end{aligned} \tag{5.1}$$

where c is a constant and the parameters $r_0, s_0, \varepsilon_0, \gamma_0, K_0$ are defined as $r, s, \varepsilon, \gamma, 1$ respectively.

For later use, we define the resonant sets useful for the part of measure estimate:

$$\mathcal{R}^\nu = \bigcup_{|k| \geq K_{\nu-1}, nm} \left(\mathcal{R}_k^\nu \cup \mathcal{R}_{k,n}^\nu \cup \mathcal{R}_{k,nm}^\nu \right) \tag{5.2}$$

where each part is defined by

$$\mathcal{R}_k^\nu = \{ \xi \in \mathcal{O}_{\nu-1} : |\langle k, \omega_\nu \rangle| < \frac{\gamma_\nu}{|k|^\tau} \} \tag{5.3}$$

$$\mathcal{R}_{k,n}^\nu = \{ \xi \in \mathcal{O}_{\nu-1} : |\det(\langle k, \omega_\nu \rangle \pm A_n^\nu \otimes I_2)| < \frac{\gamma_\nu}{|k|^\tau} \} \tag{5.4}$$

$$\mathcal{R}_{k,nm}^\nu = \{ \xi \in \mathcal{O}_{\nu-1} : |\det(\langle k, \omega_\nu \rangle \pm A_n^\nu \otimes I_2 \pm I_2 \otimes A_m^\nu)| < \frac{\gamma_\nu}{|k|^\tau} \} \tag{5.5}$$

Now we could state the iterative lemma as follows:

Lemma 5.1 *Let ε is small enough and $\nu \geq 0$, assume that we are at the ν th step.*

(1) $N_\nu + \mathcal{A}_\nu + \mathcal{B}_\nu + \bar{\mathcal{B}}_\nu$ is the normal form depending on the parameter ξ , where

$$\begin{aligned}
N_\nu &= \langle k, \omega_\nu \rangle + \sum_{n \in \mathbb{Z}_1^2} \Omega_n^\nu z_n \bar{z}_n \\
\mathcal{A}_\nu &= \sum_{n \in \mathcal{L}_1} a_n^\nu e^{i(\theta_i - \theta_j)} z_n \bar{z}_m \\
\mathcal{B}_\nu &= \sum_{n \in \mathcal{L}_2} a_n^\nu e^{-i(\theta_i + \theta_j)} z_n z_m \\
\bar{\mathcal{B}}_\nu &= \sum_{n \in \mathcal{L}_2} a_n^\nu e^{i(\theta_i + \theta_j)} \bar{z}_n \bar{z}_m
\end{aligned}$$

and satisfying the following small divisor conditions:

$$\begin{aligned} |\langle k, \omega_\nu \rangle| &\geq \frac{\gamma_\nu}{|k|^\tau} \\ |\det(\langle k, \omega_\nu \rangle + A_n^\nu \otimes I_2)| &\geq \frac{\gamma_\nu}{|k|^\tau} \\ |\det(\langle k, \omega_\nu \rangle + A_n^\nu \otimes I_2 \pm I_2 \otimes A_{n'}^\nu)| &\geq \frac{\gamma_\nu}{|k|^\tau} \end{aligned}$$

where the matrix is defined as

$$\begin{aligned} A_n^\nu &= \Omega_n \quad n \in \mathbb{Z}_1^2 \setminus (\mathcal{L}_1 \cup \mathcal{L}_2) \\ A_n^\nu &= \begin{pmatrix} \Omega_n^\nu + \omega_i^\nu & a_n^\nu \\ a_m^\nu & \Omega_m^\nu + \omega_j^\nu \end{pmatrix} \quad n \in \mathcal{L}_1 \\ A_n^\nu &= \begin{pmatrix} \Omega_n^\nu - \omega_i^\nu & -a_n^\nu \\ \bar{a}_m^\nu & \Omega_m^\nu - \omega_j^\nu \end{pmatrix} \quad n \in \mathcal{L}_2 \end{aligned}$$

and the parameter ξ is in a closed set \mathcal{O}_ν of \mathbb{R}^b .

(2) ω_ν, Ω_ν and a_n^ν are C_W^4 smooth and satisfy the condition ($\delta = \min\{\bar{a} - a, \iota\}$)

$$|\omega_{\nu-1} - \omega_\nu| < \varepsilon_{\nu-1}, |\Omega_n^\nu - \Omega_n^{\nu-1}| < \varepsilon_{\nu-1}|n|^{-\delta}, |a_n^{\nu-1} - a_n^\nu| < \varepsilon_{\nu-1}|n|^{-\delta}$$

(3) The perturbation P_ν satisfy condition (A5) and $\|X_{P_\nu}\|_{D(r_\nu, s_\nu), \mathcal{O}_\nu} < \varepsilon_\nu$.

Then there exists a closed subset $\mathcal{O}_{\nu+1} \subseteq \mathcal{O}_\nu$ defined by

$$\mathcal{O}_{\nu+1} = \mathcal{O}_\nu \setminus \mathcal{R}^\nu$$

and a symplectic transformation of variables

$$\Phi_\nu : D(r_{\nu+1}, s_{\nu+1}) \times \mathcal{O}_\nu \rightarrow D(r_\nu, s_\nu) \times \mathcal{O}_{\nu-1}$$

such that on the domain $D(r_{\nu+1}, s_{\nu+1}) \times \mathcal{O}_\nu$, the Hamiltonian has the form

$$\begin{aligned} H_{\nu+1} &= N_{\nu+1} + \mathcal{A}_{\nu+1} + \mathcal{B}_{\nu+1} + \bar{\mathcal{B}}_{\nu+1} \\ &= \langle k, \omega_{\nu+1} \rangle + \sum_{n \in \mathbb{Z}_1^2} \Omega_n^{\nu+1} z_n \bar{z}_n \\ &\quad + \sum_{n \in \mathcal{L}_1} a_n^{\nu+1} e^{i(\theta_i - \theta_j)} z_n \bar{z}_m \\ &\quad + \sum_{n \in \mathcal{L}_2} a_n^{\nu+1} e^{-i(\theta_i + \theta_j)} z_n z_m \\ &\quad + \sum_{n \in \mathcal{L}_2} a_n^{\nu+1} e^{i(\theta_i + \theta_j)} \bar{z}_n \bar{z}_m \\ &\quad + P_{\nu+1} \end{aligned}$$

with

$$|\omega_{\nu+1} - \omega_\nu| < \varepsilon_\nu, \quad |\Omega_n^{\nu+1} - \Omega_n^\nu| < \varepsilon_\nu |n|^{-\delta}, \quad |a_n^{\nu+1} - a_n^\nu| < \varepsilon_\nu |n|^{-\delta}$$

The new perturbation $P_{\nu+1}$ satisfy condition (A5) and

$$\|X_{P_{\nu+1}}\|_{D(r_{\nu+1}, s_{\nu+1}), \mathcal{O}_{\nu+1}} < \varepsilon_{\nu+1}$$

Now assume that the assumption of **(A1) – (A5)** is satisfied. We could apply the iterative lemma at the $\nu = 0$ step as long as ε_0, γ_0 are sufficiently small. By an inductive way, we get the sequence

$$\begin{aligned}\mathcal{O}_{\nu+1} &\subseteq \mathcal{O}, \\ \Phi^\nu &= \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_\nu : D(r_{\nu+1}, s_{\nu+1}) \times \mathcal{O}_\nu \rightarrow D(r_0, s_0) \times \mathcal{O} \\ H \circ \Phi^\nu &= N_{\nu+1} + \mathcal{A}_{\nu+1} + \mathcal{B}_{\nu+1} + \bar{\mathcal{B}}_{\nu+1} + P_{\nu+1}\end{aligned}$$

Let $\check{\mathcal{O}} = \bigcap_{\nu=0}^\infty \mathcal{O}_\nu$. It's easy to conclude that $N_\nu, \mathcal{A}_\nu, \mathcal{B}_\nu, \bar{\mathcal{B}}_\nu, \Phi^\nu, D\Phi^\nu$ all converge uniformly on $D(\frac{1}{2}r, 0) \times \check{\mathcal{O}}$ with

$$\begin{aligned}N_\infty &= \langle \omega_\infty, I \rangle + \sum_{n \in \mathbb{Z}_1^2} \Omega_n^\infty z_n \bar{z}_n \\ \mathcal{A}_\infty &= \sum_{n \in \mathcal{L}_1} a_n^\infty e^{i(\theta_i - \theta_j)} z_n \bar{z}_m \\ \mathcal{B}_\infty &= \sum_{n \in \mathcal{L}_2} a_n^\infty e^{-i(\theta_i + \theta_j)} z_n z_m \\ \bar{\mathcal{B}}_\infty &= \sum_{n \in \mathcal{L}_2} a_n^\infty e^{i(\theta_i + \theta_j)} \bar{z}_n \bar{z}_m\end{aligned}$$

and

$$\varepsilon_{\nu+1} = c\gamma^{-16}(r_\nu - r_{\nu+1})^{-c}\varepsilon^{\frac{4}{3}} \rightarrow 0 \quad (5.6)$$

provided that ε is sufficiently small.

Let ϕ_H^t be the Hamiltonian flow induced by X_H . By $H_{\nu+1} = H \circ \Phi^\nu$ one has

$$\phi_H^t \circ \Phi^\nu = \Phi^\nu \circ \phi_{H_{\nu+1}}^t$$

and by the uniform convergence of all the related parameters, we get

$$\phi_H^t \circ \Phi^\infty = \Phi^\infty \circ \phi_{H_\infty}^t$$

and

$$\Phi^\infty : D(\frac{1}{2}r, 0) \times \check{\mathcal{O}} \rightarrow D(r, s) \times \mathcal{O}$$

For parameters $\xi \in \check{\mathcal{O}}$, one has

$$\phi_H^t(\Phi^\infty(\mathbb{T}^b \times \{\xi\})) = \Phi^\infty \phi_{N_\infty + \mathcal{A}_\infty + \mathcal{B}_\infty + \bar{\mathcal{B}}_\infty}^t = \Phi^\infty(\mathbb{T}^b \times \{\xi\})$$

This means that $\Phi^\infty(\mathbb{T}^b \times \{\xi\})$ is an embedded torus which is invariant for the original perturbed Hamiltonian system at $\xi \in \check{\mathcal{O}}$. The frequencies $\omega_\infty(\xi)$ associated to the tori $\Phi^\infty(\mathbb{T}^b \times \{\xi\})$ is slightly different from $\omega(\xi)$. The normal behavior of the invariant torus is governed by normal frequencies Ω_n^∞ . \square

6 Measure Estimate

Recall the resonant sets at the ν th step $\mathcal{R}^\nu = \bigcup_{|k| \geq K_{\nu-1}, nm} (\mathcal{R}_k^\nu \cup \mathcal{R}_{k,n}^\nu \cup \mathcal{R}_{k,nm}^\nu)$. To estimate its measure, we need to estimate each single set $\mathcal{R}_k^\nu, \mathcal{R}_{k,n}^\nu, \mathcal{R}_{k,nm}^\nu$ first.

Lemma 6.1 Fix $|k| \geq K_{\nu-1}, n, m$, one has

$$meas \left(\mathcal{R}_k^\nu \cup \mathcal{R}_{k,n}^\nu \cup \mathcal{R}_{k,nm}^\nu \right) < c \frac{\gamma_\nu^{\frac{1}{4}}}{|k|^{\frac{1}{4}\tau}}$$

Proof: One has that $\omega_\nu(\xi) = \omega(\xi) + \sum_{j=0}^{\nu-1} P_{0l00}^j(\xi)$ with $\sum_{0 \leq j \leq \nu-1} |P_{0l00}^j(\xi)|_{\mathcal{O}_{\nu-1}} < \varepsilon$, and $\Omega_n^\nu(\xi) = \Omega_n(\xi) + \sum_{0 \leq j \leq \nu-1} P_{nn}^{011,j}(\xi)$ with $\sum_{0 \leq j \leq \nu-1} |P_{nn}^{011,j}(\xi)|_{\mathcal{O}_{\nu-1}} < \frac{\varepsilon}{|n|^\delta}$ ($\delta = \min\{\bar{a} - a, \iota\} > 0$) Similar results also hold for a_n . So it's easy to conclude that

$$\max_{1 \leq j \leq 4} |\partial_\xi^j \det(\langle k, \omega_\nu \rangle \pm A_n \otimes I_2 \pm I_2 \otimes A_m)| \geq c|k|$$

Then the result is obvious. \square

Lemma 6.2 The whole measure we need to exclude during the KAM procedure is

$$meas \left(\bigcup_{\nu \geq 0} \mathcal{R}^\nu \right) < c\gamma^\varsigma \quad \varsigma > 0$$

Proof: Fix one ν and one k , and we only estimate the most complicated term:

$$\bigcup_{n,m} \left\{ \xi \in \mathcal{O}_{\nu-1} : |\det(\langle k, \omega_\nu \rangle + A_n \otimes I_2 - I_2 \otimes A_m)| < \frac{\gamma_\nu}{|k|^\tau} \right\}$$

Consider its diagonal entry, we only consider one element. If $|n|^2 - |m|^2 = l \geq c|k|$, then $\mathcal{R}_{k,nm}^\nu = \emptyset$ otherwise we assume $|n| \geq |m|$, by the regularity property, we get

$$|\Omega_n^\nu - \Omega_m^\nu - \varepsilon_0^{-4}l| < O(|m|^{-\delta}) \quad \delta = \min\{\bar{a} - a, \iota\} > 0$$

so we have that

$$\mathcal{R}_{k,nm}^\nu \subseteq \mathcal{Q}_{k,lm}^\nu = \left\{ \xi : |\det(\langle k, \omega_\nu \rangle + A_n \otimes I_2 - I_2 \otimes A_m)| < \frac{\gamma_\nu}{|k|^\tau} + O(|m|^{-\delta}) \right\}$$

it's easy to see that $\mathcal{Q}_{lm}^\nu \subseteq \mathcal{Q}_{lm_0}^\nu$ for $|m| \geq |m_0|$. By Lemma 6.1 we have

$$\begin{aligned} meas \left(\bigcup_{|l| \leq c|k|} \bigcup_{|n|^2 - |m|^2 = l} \mathcal{R}_{k,nm}^\nu \right) &\leq \sum_{|l| \leq c|k|} \sum_{|m| \leq |m_0|} meas(\mathcal{R}_{k,nm}^\nu) + \sum_{l \leq c|k|} meas(\mathcal{Q}_{k,lm_0}^\nu) \\ &\leq c \left(\frac{|k| \gamma^{\frac{1}{4}} |m_0|^2}{|k|^{\frac{\tau}{4}}} + O(|k| |m_0|^{-\delta}) \right) \end{aligned}$$

By choosing appropriate m_0 to reach $\frac{\gamma^{\frac{1}{4}} |m_0|^2}{|k|^{\frac{\tau}{4}}} = |m_0|^{-\delta}$ (just let $|m_0| = \left(\frac{|k|^\tau}{\gamma} \right)^{\frac{1}{4\delta+8}}$). Then we get the estimate

$$meas \left(\bigcup_{|l| \leq c|k|} \bigcup_{|n|^2 - |m|^2 = l} \mathcal{R}_{k,nm}^\nu \right) < c \frac{\gamma^{\frac{\delta}{4\delta+8}}}{|k|^{\frac{\delta\tau}{4\delta+8}-1}}$$

justify the parameter τ appropriately and we get the result. \square

7 Appendix

In this part we give a precise method to construct the admissible set $S = \{i_1 = (x_1, y_1), i_2 = (x_2, y_2), \dots, i_b = (x_b, y_b)\}$. It's modified from the appendix in [11] and we omit some detailed calculation which has been done in [11]. The points in S will be defined in an inductive way. The first point $(x_1, y_1) \in \mathbb{Z}_{odd}^2$ is chosen as $x_1 > b^2, y_1 = 2x_1^{5^b}$ and the second $x_2 = x_1^5, y_2 = 2x_2^{5^b}$. If we have chosen the first j points i_1, i_2, \dots, i_j , then we define

$$x_{j+1} = x_j^5 \left(\prod_{1 \leq l < m \leq j} ((x_m - x_l)^2 + (y_m - y_l)^2) + 1 \right) \quad 2 \leq j \leq b-1$$

$$y_{j+1} = 2x_{j+1}^{5^b} \quad 2 \leq j \leq b-1$$

Recall the condition of admissible set (Proposition 1). We verify the conditions one by one. Given three points $c, d, f \in S$, it's easy to see that

$$\langle c - d, d - f \rangle = (c_1 - d_1)(c_2 - d_2) + (d_1 - f_1)(d_2 - f_2) > 0$$

So any three points in S can't be three vertices of a rectangle.

To verify condition ②-④, following the appendix in [11], it suffices to prove that each equation set in the following has no integer solution in \mathbb{Z}_1^2 for $c, d, f, g \in S$ and $\{c, d\} \neq \{f, g\}$.

$$\begin{cases} \langle n - g, g - f \rangle = 0 \\ \langle n - c, c - d \rangle = 0 \end{cases}$$

$$\begin{cases} \langle n - g, n - f \rangle = 0 \\ \langle n - c, n - d \rangle = 0 \end{cases}$$

$$\begin{cases} \langle n - g, g - f \rangle = 0 \\ \langle n - c, n - d \rangle = 0 \end{cases}$$

The three equation sets correspond to condition ②, ③, ④ respectively and we denote them by **I**, **II**, **III** respectively.

For **I**,

I - i: If only one element of $\{|c|, |d|, |f|, |g|\}$ reaches the maximum value of them.

I - i - (1): $|d|$ or $|f|$ reaches the maximum. Without loss of generality, just assume d . It's easy to get that

$$n_2 = c_2 + \frac{(g_1 - c_1)(g_1 - f_1)(c_1 - d_1) + (g_2 - c_2)(c_1 - d_1)(g_2 - f_2)}{(c_1 - d_1)(g_2 - f_2) - (c_2 - d_2)(g_1 - f_1)}$$

By the fact $d_2 \gg$ other ones, the second term in the right side couldn't be an integer, implying $n_2 \notin \mathbb{Z}$, which is a contradiction.

I – i – (2) : $|g|$ or $|c|$ reaches the maximum. Without lose of generality, just assume $|g|$. By calculation, we get

$$\begin{aligned} n_2 &= g_2 + 2(c_1^{5^b-1} + c_1^{5^b-2}d_1 + \cdots + c_1d_1^{5^b-2} + d_1^{5^b-1})(g_1 - f_1) \\ &+ \frac{g_1 - c_1 + 4(g_1 - f_1)A^2 + 4(f_1^{5^b} - c_1^{5^b})A}{2(g_1^{5^b-1} + g_1^{5^b-2}f_1 + \cdots + g_1f_1^{5^b-2} + f_1^{5^b-1}) - 2A} \end{aligned}$$

where $A = (c_1^{5^b-1} + c_1^{5^b-2}d_1 + \cdots + c_1d_1^{5^b-2} + d_1^{5^b-1})$. By the fact $g_1 \gg$ other ones, we know that the last term is not an integer, so $n_2 \notin \mathbb{Z}$, which is a contradiction.

I – ii : If two elements of $|c|, |d|, |f|, |g|$ reach the maximum, by the structure of S we know that these two points must be equal, and the case when three points reach the maximum will not happen.

I – ii – (1) : when $|d| = |g|$ reach the maximum, we have $d = g$. By calculation we get

$$n_2 = \frac{4g_1^{5^b}A - 4c_1^{5^b}B + (g_1 - c_1)}{4(A - B)}$$

where $A = (g_1^{5^b-1} + g_1^{5^b-2}f_1 + \cdots + g_1f_1^{5^b-2} + f_1^{5^b-1})$ and $B = (g_1^{5^b-1} + g_1^{5^b-2}c_1 + \cdots + g_1c_1^{5^b-2} + c_1^{5^b-1})$ without losing generality, we assume $|c| < |f|$ (the case $|c| = |f|$ will not happen). According to the structure of S, it's easy to see that in the expression above, the denominator is divisible by $|c_1|^4$, in the numerator, all terms are divisible by $|c_1|^4$ except for c_1 . It means that $n_2 \notin \mathbb{Z}$ which is a contradiction.

I – ii – (2) : If $|d| = |f|$ reach the maximum, we have $d = f$, then we have

$$\langle n - g, g - f \rangle = \langle n - g, g - d \rangle = \langle n - c, c - f \rangle = 0$$

In this case c, d, g are three vertices of a rectangular, which is a contradiction.

I – ii – (3) : If $|c| = |g|$ reach the maximum, we have $c = g$. From

$$\langle n - g, g - f \rangle = \langle n - g, g - d \rangle = \langle n - c, c - f \rangle = 0$$

we know d, f, g lie on the same line, which is a contradiction.

Now we turn to **III**.

III – i : If only one element of $|c|, |d|, |f|, |g|$ reaches the maximum value of them and each one is different from others.

III – i – (1) : $|d|$ reaches the maximum.

We have

$$\langle n - c, n - d \rangle = 0 \quad \langle n - g, g - f \rangle = 0$$

We take g as the origin and from above we get an equation about n_1 .

$$(f_1^2 + f_2^2)n_1^2 + ((c_2 + d_2)f_1f_2 - (c_1 + d_1)f_2^2)n_1 + f_2^2(c_1d_1 + c_2d_2) = 0$$

The discriminant of the equation is

$$\Delta = \left(f_1f_2d_2 + \left(c_2f_1f_2 - c_1f_2^2 - d_1f_2^2 - \frac{2c_2f_2(f_1^2 + f_2^2)}{f_1} \right) - \alpha \right)^2$$

where $\alpha \sim \frac{d_1}{d_2} \ll \frac{1}{|f_1|}$ so we get

$$n_1 = \frac{-((c_2 + d_2)f_1f_2 - (c_1 + d_1)f_2^2) \pm \sqrt{\Delta}}{2(f_1^2 + f_2^2)}$$

while $\sqrt{\Delta} = \text{some integer} + \alpha$, so the numerator in the above expression is not an integer. So we conclude that $n_1 \notin \mathbb{Z}$, which is a contradiction.

III – i – (2) $|f|$ reaches the maximum.

As before, we take g as the origin and we get

$$(f_1^2 + f_2^2)n_1^2 + ((c_2 + d_2)f_1f_2 - (c_1 + d_1)f_2^2)n_1 + (c_1d_1 + c_2d_2)f_2^2 = 0$$

The discriminant is

$$\Delta = (c_1 - d_1)^2 f_2^4 - 4c_2d_2f_2^4 - 2(c_1 + d_1)(c_2 + d_2)f_1f_2^3 + (c_2 - d_2)^2 f_1^2 f_2^2 - 4c_1d_1f_1^2 f_2^2 < 0$$

by the fact $c_2d_2 \gg (c_1 - d_1)^2$ and $f_2 \gg$ other terms.

III – i – (3) $|g|$ reaches the maximum.

It's easy to see $|n|^2 > |g|^2 - |f|^2$ and we get

$$|n|^2 + |m|^2 - |c|^2 - |d|^2 > 0$$

III – ii If only one element of $\{|c|, |d|, |f|, |g|\}$ reaches the maximum and two of the others are the same.

III – ii – (1) $|d|$ reaches the maximum, and $c = g$, without losing generality, we just assume $c = g$. We could assume $|f| > |c|$, the case $|f| < |c|$ is similar.

In this case, by calculation we get

$$n_2 = c_2 + \frac{(f_1 - c_1)^2 d_2 - (f_1 - c_1)(f_2 - c_2)d_1}{(f_1 - c_1)^2 + (f_2 - c_2)^2} + \frac{2c_1f_1(f_1^{5b-2} + f_1^{5b-3}c_1 + \dots + f_1c_1^{5b-3} + c_1^{5b-2})}{1 + 4(f_1^{5b-1} + f_1^{5b-2}c_1 + \dots + f_1c_1^{5b-2} + c_1^{5b-1})^2}$$

By the structure of S , we could conclude that the term

$$\frac{(f_1 - c_1)^2 d_2 - (f_2 - c_2)(f_1 - c_1)d_1}{(f_1 - c_1)^2 + (f_2 - c_2)^2}$$

must has a form of an integer plus $\frac{2(f_1 - c_1)^2 f_1^\kappa - (f_2 - c_2)(f_1 - c_1)f_1^{\tilde{\kappa}}}{(f_1 - c_1)^2 + (f_2 - c_2)^2}$ where $0 < \kappa \leq 5^b$ and $5 \leq \tilde{\kappa} \leq 5^{b-1}$. Then we have $\frac{(f_1 - c_1)^2 f_1^\kappa}{(f_1 - c_1)^2 + (f_2 - c_2)^2} \leq \frac{1}{f_1^{5b-2}}$ and $\frac{c_1f_1(f_1^{5b-2} + f_1^{5b-3}c_1 + \dots + f_1c_1^{5b-3} + c_1^{5b-2})}{1 + (f_1^{5b-1} + f_1^{5b-2}c_1 + \dots + f_1c_1^{5b-2} + c_1^{5b-1})^2} \leq \frac{1}{f_1^{5b-2}}$ while $\frac{(f_2 - c_2)(f_1 - c_1)f_1^{\tilde{\kappa}}}{(f_2 - c_2)^2 + (f_1 - c_1)^2} > f_1^{6-5^b}$ which is much larger than the former two, and these three terms all $\ll 1$, hence $n_2 \notin \mathbb{Z}$. The case $c = f$ is similar to **III – i – (1)**.

III – ii – (2) $|f|$ reaches the maximum and $g = c$ or $g = d$. Without losing generality, we just assume $g = c$.

We have

$$\langle n - c, n - d \rangle = 0 \quad \langle n - c, c - f \rangle = 0$$

If we take c as the origin, we get $n_2 = \frac{f_1^2 d_2 - f_1 f_2 d_1}{f_1^2 + f_2^2} \notin \mathbb{Z}$

III – ii – (3) $|g|$ reaches the maximum. It's similar to **III – i – (3)**.

III – iii If two elements of $\{|c|, |d|, |f|, |g|\}$ reach the maximum of their values.

III – iii – (1) $|g| = |d|$ reach the maximum. We have $g = d$. Then we get

$$n_1 = c_1 + d_1 - f_1 + \frac{(d_1 - f_1)^2(d_1 - c_1) + (f_2 - c_2)(d_1 - f_1)(d_2 - f_2) - (d_1 - f_1)^3}{(d_1 - f_1)^2 + (d_2 - f_2)^2}$$

which implies that $n_1 \notin \mathbb{Z}$ due to the fact $d_2 \gg$ other terms.

III – iii – (2) $|f| = |d|$ reach the maximum. We have $f = d$ and take g as the origin as before. Then we get

$$(d_1^2 + d_2^2)n_1^2 + (d_1d_2c_2 - c_1d_2^2)n_1 + c_1d_1d_2^2 + c_2d_2^3 = 0$$

The discriminant is

$$\Delta = (d_1d_2c_2 - c_1d_2^2)^2 - 4(d_1^2 + d_2^2)(c_1d_1d_2^2 + c_2d_2^3) < 0$$

due to the fact $d_2 \gg$ other terms. So n_1 doesn't exist, which is a contradiction.

At last we turn to **II**.

We know $\{c, d\} \notin \{f, g\}$. If $\{c, d\} \cap \{f, g\} \neq \emptyset$, without losing generality we just assume $c = f$. The equation becomes

$$\langle n - g, g - d \rangle = 0 \quad \langle n - c, n - d \rangle = 0$$

We have proved that it has no solution in \mathbb{Z}^2 in **III – (iii) – (2)**.

Now we concentrate on the remaining case when the four elements are different from each other. Without losing generality, we just assume $|d|$ reach the maximum of their values. It's easy to see $|n|^2 \ll d_1$ and

$$n_2 = c_2 + \frac{(c_1 - n_1)d_1 - c_2(c_2 - f_2 - g_2) - \langle f, g \rangle - n_1(c_1 - f_1 - g_1)}{c_2 + d_2 - f_2 - g_2}$$

If $c_1 = n_1$, then in the above expression the numerator is smaller than d_1 , and if $c_1 \neq n_1$, the numerator is smaller than $\frac{d_2}{2}$, we still have $n_2 \notin \mathbb{Z}$, which is a contradiction.

References

- [1] M.Berti, P.Bolle. Sobolev quasi periodic solutions of multidimensional wave equations with a multiplicative potential. *Nonlinearity* **25** (2012), 2579-2613.
- [2] M.Berti, P.Bolle. Quasi-periodic solutions with Sobolev regularity of NLS on \mathbb{T}^d and a multiplicative potential. *J.European Math.Society*, **15** (2013) 229-286.
- [3] J.Bourgain. Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equation. *Ann. Math.* **148** (1998) 363-439.
- [4] J.Bourgain. Green's Function Estimates for Lattice Schrödinger Operators and Applications. *Ann. of Math.Stud.* vol.158, Princeton University Press, Princeton, NJ, 2005.
- [5] L.Chierchia, J.You. KAM tori for 1D nonlinear wave equations with periodic boundary conditions. *Comm.Math. Phys.* **211** (2000) 498-525
- [6] W.Craig, C.E.Wayne. Newton's method and periodic solutions of nonlinear wave equations. *Comm. Pure Appl. Math.* **46** (1993) 1409-1498.
- [7] L.H.Eliasson, B.Grebert, S.B.Kuksin: KAM for the nonlinear beam equation. *preprint*
- [8] L.H.Eliasson, S.B.Kuksin: KAM for nonlinear Schrödinger equation. *Ann.of Math.* **172** (2010) 371 – 435.
- [9] L.H.Eliasson, S.B.Kuksin. Infinite Töeplitz-Lipschitz matrices and operators. *Z.Angew.Math.Phys.* **59** (2008) 24-50.
- [10] L.H.Eliasson, S.B.Kuksin. On reducibility of Schrödinger equations with quasiperiodic in time potentials. *Comm.Math.Phys.* **286** (2009) 125-135.
- [11] J.Geng, X.Xu, J.You. An infinite dimensional KAM theorem and its application to the two dimensional cubic Schrödinger equation. *Adv.Math.* **226** (2011) 5361 – 5402
- [12] J.Geng, Y.Yi. Quasi-periodic solutions in a nonlinear Schrödinger equation. *J.Differential Equations* **233** (2007) 512-542.
- [13] J.Geng, J.You. A KAM theorem for one dimensional Schrödinger equation with periodic boundary conditions. *J.Differential Equations* **209** (2005) 1-56.
- [14] J.Geng, J.You. A KAM theorem for Hamiltonian partial differential equations in higher dimensional spaces. *Comm. Math. Phys.* **262** (2006) 343-372.
- [15] J.Geng, J.You. KAM tori for higher dimensional beam equations with constant potentials. *Nonlinearity* **19** (2006) 2405-2423.
- [16] B.Grebert, E. Paturel. KAM for the Klein-Gordon equation on S^d . *Boll. Unione. Mat. Ital.* **9** (2016) 237-288.

- [17] B.Grebart, L. Thomann. KAM for the quantum harmonic oscillator. *Comm. Math. Phys.* **307** (2011) 383-427.
- [18] T.Kappeler, J.Pöschel. *KdV & KAM*, Springer, 2003.
- [19] S.B.Kuksin. Hamiltonian perturbations of infinite-dimensional linear systems with an imaginary spectrum. *Func.Anal.Appl.* **21** (1987) 192-205.
- [20] S.B.Kuksin. *Nearly Integrable Infinite Dimensional Hamiltonian Systems*. Lecture Notes in Math. vol. 1556, Springer, Berlin, 1992.
- [21] S.B.Kuksin, J.Pöschel. Invariant Cantor manifolds of quasiperiodic oscillations for a nonlinear Schrödinger equation. *Ann. of Math.* **143** (1996) 149-179.
- [22] S.B.Kuksin. *Analysis of Hamiltonian PDEs*. Oxford University Press. 2000.
- [23] J.Pöschel. On elliptic lower dimensional tori in Hamiltonian systems. *Math. Z.* **202** (1989) 559-608.
- [24] J.Pöschel. A KAM theorem for some nonlinear partial differential equations. *Ann. Sc Norm. Sup. Pisa Cl. Sci.* **23** (1996) 119-148.
- [25] J.Pöschel. Quasi-periodic solutions for a nonlinear wave equation. *Comment. Math. Helv.* **71** (1996) 269-296.
- [26] C.Procesi, M.Procesi. A normal form for the Schrödinger equation with analytic non-linearities. *Comm. Math. Phys.* **312** (2012) 501-557.
- [27] C.Procesi, M.Procesi. A KAM algorithm for the resonant non-linear Schrödinger equation. *Adv.Math.* **272** (2015) 399-470.
- [28] C.Procesi, M.Procesi. Reducible quasi-periodic solutions for the nonlinear Schrödinger equation. *Boll. Unione. Mat. Ital.* **9** (2016) 189-236.
- [29] M.Procesi. A normal form for beam and non-local nonlinear Schrödinger equations. *J.Phys.A.* **43** (2010) 434028, 13 pp.
- [30] W.M.Wang. Energy supercritical nonlinear Schrödinger equations: quasiperiodic solutions. *Duke Math.J.* **165** (2016) 1129-1192.
- [31] C.E.Wayne. Periodic and quasi-periodic solutions for nonlinear wave equations via KAM theory. *Com. Math. Phys.* **127** (1990) 479-528.
- [32] X.Yuan. Quasi-periodic solutions of completely resonant nonlinear wave equations. *J.Differential Equations.* **230** (2006) 213-274.